

- load tất cả các file CSV trong folder (assign nó vào 1 tên)
 - `pd.concat()` dùng để gộp nhiều DataFrame lại với nhau
 - money data: số (`astype, pd_to_numeric`) , replace \$ -> khoảng trống
 - dùng split để tách ra (thành 2 cột riêng) -> assign name từng cái
- GenderPop 2341093M_2489527F

Men	Woman
-----	-------

2341093M	2489527F
----------	----------

- thay M -> khoảng trống (xóa M) -> convert qua số
- giải quyết NaN thì mình fill na bằng cách
(Total Population - Men)
(60% bỏ luôn cột đó , dưới 5% thì drop NaN)